21PRTS

09/254344
300 Rec'd PCT/PTO 05 MAR 1999

Specification

RNA Polymerase

Technical Field

The present invention relates to mutant RNA polymerases useful for methods for determining nucleotide sequence of DNA and the like.

Background Art

The polymerase chain reaction (PCR) method is an excellent method, and its utilization has expanded year by year [Randall K. Saiki et al. (1988) Science 239, 487-491]. In the PCR method, even one molecule of DNA fragment can be amplified. The method for sequencing PCR amplified products without cloning them (the direct sequencing method) is also a useful method [Corinne Wong et al. (1988) Nature, 330, 384-386]. This technique does not require construction of libraries and screening of such libraries, and is a quick method capable of simultaneously obtaining sequence information of many samples.

However, the above direct sequencing method suffers from two major problems.

One is that primers and 2'-deoxyribonucleoside 5'-triphosphates (2'-dNTPs) not incorporated remain in a reaction system, and the remained substances inhibit sequencing reactions. Therefore, in conventional methods, such primers and 2'-dNTPs must be removed from PCR products before sequencing. There are many methods for purification of PCR products and examples include purification by electrophoresis, ethanol precipitation, gel filtration and HPLC purification [see, for example, Dorit R.L et al. (1991) Current Protocols in Molecular Biology, Vol. 11, John Wiley and Sons, New York, 15.2.1-15.2.11]. However, these methods are complicated without exception.

The second problem is quick renaturation of PCR products.

When the PCR products are renatured into a double-stranded DNA, they are no longer single-stranded templates, and annealing between primers and single-stranded templates is inhibited. As methods for minimizing the renaturation, quenching after denaturation, biotilation of one primer and absorption of PCR products onto streptavidin-coated articles, use of exonuclease, asymmetric PCR and the like have been reported. See, for example, Barbara Bachmann et al., 1990, Nucleic Acid Res.,18, 1309-. However, most of these methods are time-consuming and very laborious.

Therefore, the present inventors proposed an absolutely novel method for determining nucleotide sequence of DNA for solving these problems. Which does not require removal of unreacted primers and 2'-deoxyribonucleoside 5'-triphosphates (2'-dNTPs) remaining in the PCR reaction system, and does not require denaturation at all. This method enables to eliminate the problem of quick renaturation of PCR reaction products [WO96/14434]. This method is a direct transcriptional sequencing method utilizing an RNA polymerase such as T7 RNA polymerase and a terminator for RNA transcription reaction (for example, 3'-deoxyribonucleoside 5'-triphosphates, 3'-dNTPs). According to this method, nucleotide sequences of DNA products amplified by the polymerase chain reaction can be used as they are for sequencing without removing primers and 2'-deoxyribonucleoside 5'-triphosphates (2'-dNTPs). In addition, because it does not require denaturation itself at all, it can avoid the problem of quick renaturation of PCR products, and hence is an extremely excellent method.

However, the present inventors further studied the above method, and found that it has a problem to be solved in order to obtain more accurate nucleotide sequence data.

In the above nucleotide sequence determination method, an RNA polymerase such as T7 RNA polymerase is used for the reaction in a mixture comprising ribonucleoside 5'-triphosphates including ATP, GTP, CTP, UTP and derivatives

thereof, and at least one 3'-deoxyribonucleotide such as 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof. In this reaction, polyribonucleotides are synthesized by sequential incorporation of ribonucleotides and deoxyribonucleotides into a ribonucleotide sequence in a manner corresponding to the sequence of templates.

However, it was found that 3'-deoxyribonucleotides and derivative thereof are unlikely to be incorporated into the sequence rather than corresponding ribonucleotides, and the occurrence of the incorporation may also vary among the ribonucleotides and the 3'-deoxyribonucleotides depending on a base group each nucleotide has. Such biased incorporation between ribonucleotides and 3'-deoxyribonucleotides, as well as among ribonucleotides having different base groups and among deoxyribonucleotides having different base groups may result in short transcription products and fluctuation of signals from labeled ribonucleotides. Therefore, it is difficult to obtain accurate sequence data even though transcription products can be obtained.

Therefore, an object of the present invention is to provide an RNA polymerase exhibiting incorporation ability with no or little bias resulting from differences in nucleotides.

In the description of the present invention, amino acid residues are represented by the conventionally used one-letter codes. For clarification, they are specifically mentioned for only those amino acids appeared in this text as follows: phenylalanine (F), tyrosine (Y), proline (P), leucine (L), and histidine (H). A numeral accompanied by the codes is a number counted from N-terminus of polymerase. For example, "F667" means that the 667th amino acid residue of this polymerase is F, and "F667Y" means that Y was substituted for F of the 667th residue.

By the way, DNA polymerases are also known to show biased incorporation resulting from difference in a base group each nucleotide has, and mutant DNA polymerases free from such

biased incorporation have also been known [Japanese Patent Unexamined Publication (KOKAI) No. (Hei) 8-205874/1996; and Proc. Natl. Acid. Sci. USA, 92:6339-6345, (1995)].

In the aforementioned literatures, it is described as follows. In the sequencing reaction utilizing T7 DNA polymerase, the 526th amino acid in the polymerase contributes to equalize nucleotide incorporation. And due to homology between T7 DNA polymerase and other DNA polymerases, the bias of incorporation of the other DNA polymerases may be reduced by replacing an amino acid residue present in their region homologous to the 526th amino acid including region in the T7 DNA polymerase. That is, Y (tyrosine) 526 of T7 DNA polymerase results in the reduced bias of efficiency for incorporation of 2'-dNTPs and 2',3'-ddNTPs. F (phenylalanine) 762 of E. coli DNA polymerase I and F (phenylalanine) 667 of Thermus aquaticus DNA polymerase (generally called Taq DNA polymerase) are the amino acid residues corresponding to Y526 of T7 DNA polymerase and the bias of these polymerases may be reduced by substituting F762Y (tyrosine) and F667Y (tyrosine) respectively for these residues.

Further, it is also described that it was suggested that modification of a region of T7 RNA polymerase corresponding to the region discussed for DNA polymerases, i.e., the residues 631-640, may change its specificity for dNTPs.

However, RNA polymerases have not been used for sequencing methods so far, and therefore the different efficiency of ribonucleotide incorporation itself has not become a problem. Under such circumstances, any mutant RNA polymerases free from the biased incorporation have of course not been known. In fact, the aforementioned Japanese Patent Unexamined Publication (KOKAI) No. (Hei) 8-205874/1996 does not mention any specific examples of modification of T7 RNA polymerase.

The region of T7 RNA polymerase mentioned above is considered to correspond to the region consisting of 9-10 amino acid residues between amino acids K and YG in the motif B

mentioned in Protein Engineering, 3:461-467, 1990, which region is particularly conserved in DNA polymerase α and I, and DNA-dependent RNA polymerases (T7 RNA polymerase is classified in these polymerases). F (phenylalanine) of the amino acid residue 762 in E. coli DNA polymerase and the amino acid residue 667 in Taq DNA polymerase, previously discussed for DNA polymerases, are observed in many of DNA polymerases classified in the type I. However, it was surprisingly found that T7 RNA polymerase does not have F (phenylalanine) in the residues 631-640 corresponding to the aforementioned region, though T7 RNA polymerase is highly homologous to DNA polymerases. Therefore, the teachings of the aforementioned literatures could not be realized as described.

Further, the present inventors attempted modification of amino acids of T7 RNA polymerase in the region corresponding to the helix O of the finger subdomain of E. coli DNA polymerase I, in which F762 of E. coli DNA polymerase I presents. However, F (phenylalanine) was not found also in the helix Z in T7 RNA polymerase, which is indicated in the steric structure reported in the literature of Sousa et al. (Nature, 364:593-599, 1993) and corresponds to the helix O of E. coli DNA polymerase I.

Under the circumstances, the present inventors originally searched for a novel RNA polymerase in order to provide an RNA polymerase which exhibits little or no bias for the incorporating ability valuable due to the kind of ribonucleotides and 3'-deoxyribonucleotides. As a result, the present invention was completed based on the findings that an RNA polymerase having an increased ability of incorporating 3'-deoxyribonucleotides and derivatives thereof can be obtained by partially modifying amino acids in a wild type RNA polymerase.

While it will be apparent from the descriptions hereinafter, the RNA polymerase of the present invention, or in particular the location of the amino acid modification thereof is not suggested nor taught at all in Japanese Patent Unexamined Publication (KOKAI) No. (Hei) 8-205874/1996, and

it was absolutely originally found by the present inventors.

Summary of the Invention

The present invention relates to an RNA polymerase consisting of a wild type RNA polymerase provided that at least one of amino acids in the wild type RNA polymerase was modified so as to enhance its ability for incorporating 3'-deoxyribonucleotides and derivatives thereof in comparison with the corresponding wild type RNA polymerase.

Brief Description of the Drawings

Figure 1 shows T7 RNA polymerase gene on the T7 phage genome and the amino acid sequence of the encoded T7 RNA polymerase (first half). The nucleotide sequence is shown in the upper sections, and the corresponding amino acid sequence is in the lower sections. The numerals for the nucleotide sequence at the right end indicate numbers of T7 phage genome registered at the DNA sequence database GeneBank (Locus T7CG, 39,937 base pairs), and the numerals of amino acids are appended from the first M (methionine) of T7 RNA polymerase starting with 1, and indicate that the full length is composed of 883 amino acid residues.

Figure 2 shows T7 RNA polymerase gene on the T7 phage genome and the amino acid sequence of the encoded T7 RNA polymerase (latter half). The nucleotide sequence is shown in the upper sections, and the corresponding amino acid sequence is in the lower sections. The numerals for the nucleotide sequence at the right end indicate numbers of T7 phage genome registered at the DNA sequence database GeneBank (Locus T7CG, 39,937 base pairs), and the numerals of amino acids are appended from the first M (methionine) of T7 RNA polymerase starting with 1, and indicate that the full length is composed of 883 amino acid residues.

Figure 3 shows alignment of amino acid sequences of the currently reported phage-derived RNA polymerases (first half). The T7 RNA polymerase at the top is used as a standard, and

the symbols . (dot) indicate the same amino acid residues as the T7 RNA polymerase, - indicates absence, and * at the bottom indicates an amino acid residue common to all of the polymerases.

Figure 4 shows alignment of amino acid sequences of the currently reported phage-derived RNA polymerases (latter half). The T7 RNA polymerase at the top is used as a standard, and the symbols . (dot) indicate the same amino acid residues as the T7 RNA polymerase, - indicates absence, and * at the bottom indicates an amino acid residue common to all of the polymerases.

Figure 5 shows details of mutated sites of T7 RNA polymerase. The outline characters indicate mutated amino acids.

Figure 6 shows alignment of amino acid sequences of T7 RNA polymerase and T3 RNA polymerase (first half). The T7 RNA polymerase at the top is used as a standard, and the symbols . (dot) indicate the same amino acid residues as the T7 RNA polymerase, - indicates absence, and * at the bottom indicates amino acid residues common to the both polymerases.

Figure 7 shows alignment of amino acid sequences of T7 RNA polymerase and T3 RNA polymerase (latter half). The T7 RNA polymerase at the top is used as a standard, and the symbols . (dot) indicate the same amino acid residues as the T7 RNA polymerase, - indicates absence, and * at the bottom indicates amino acid residues common to the both polymerases.

Figure 8 shows the sequences around the residues 641-667 of T7 RNA polymerase, and amino acid sequences of the corresponding regions of T3 RNA polymerase, K11 RNA polymerase and SP6 RNA polymerase. While all of the residues are shown for T7RNA polymerase, the corresponding residues are indicated with . (dot) for T3, K11, and SP6 when they are the same as those of T7.

Figure 9 shows a construction map of pT7R, a plasmid expressing wild type T7 RNA polymerase.

Figure 10 shows a construction map of pT7RF644Y, a

plasmid expressing a mutant T7 RNA polymerase F644Y.

Figure 11 shows a construction map of an improved version of plasmid pT7R, pT7R-Xho, having a restriction endonuclease XhoI site in the T7 RNA polymerase gene.

Figure 12 shows a construction map of pT7RL665P/F667Y, a plasmid expressing a mutant T7RNA polymerase L665P/F667Y.

Figure 13 demonstrates improvement of incorporation rate of dye terminator by mutant T7 RNA polymerases. The results of wild type T7 RNA polymerase (WT), mutant T7 RNA polymerase F644Y (F644Y), and mutant T7 RNA polymerase L665P/F667Y (F667Y) are shown.

Figure 14 demonstrates improvement of incorporation rate of dye terminator by mutant T7 RNA polymerase F644Y. The results of wild type T7 RNA polymerase (WT), and mutant T7 RNA polymerase F644Y(F644Y) are indicated as an electropherogram.

Figure 15 demonstrates improvement of incorporation rate of dye terminator by mutant T7 RNA polymerase L665P/F667Y. The results of wild type T7 RNA polymerase (WT), and mutant T7 RNA polymerase L665P/F667Y (F667Y) are indicated as an electropherogram.

Figure 16 shows an example of sequencing reaction. The reaction was performed by using wild type T7 RNA polymerase (WT), mutant T7 RNA polymerase F644Y (F644Y), or a mutant T7 RNA polymerase L665P/F667Y (F667Y). Sequencing patterns of the same area are shown, and it can be observed that the sequencing could not be correctly performed in the wild type T7 RNA polymerase (WT) (top), because the base call did not correctly function, and interval of bases became too narrow (representations of the bases overlap).

Figure 17 shows a construction map of pT7R F644Y/L665P/F667Y, a plasmid expressing a mutant T7RNA polymerase F644Y/L665P/F667Y.

Figure 18 (1)-(4) demonstrate improvement of incorporation rate of dye terminator by mutant T7 RNA polymerase F644Y/L665P/F667Y as an electropherogram.

Embodiments for Carrying Out the Invention

According to the present invention, the "wild type RNA polymerase" include any naturally occurring RNA polymerases. In addition, the "wild type RNA polymerase" may be a wild type RNA polymerase having substitution, insertion and/or deletion of amino acids which are not the modification for obtaining increased ability for incorporating 3'-deoxyribonucleotide and derivatives thereof in comparison with the corresponding wild type RNA polymerase. That is, wild type RNA polymerases artificially modified with a purpose other than that described above are included in the above "wild type RNA polymerase". However, it is suitable to make such substitution, insertion and/or deletion of amino acids to the extent that the activity of RNA polymerase is maintained.

Examples of the "wild type RNA polymerase" include RNA polymerases derived from T7 phage, T3 phage, SP6 phage, K11 phage and the like. However, it is not limited to these RNA polymerases.

The "wild type RNA polymerase" according to the present invention include naturally occurring thermostable RNA polymerases, and naturally occurring RNA polymerases artificially modified (i.e. having substitution, insertion and/or deletion of amino acids) in order to impart thermostablity. However, it is suitable to make the modification for imparting thermostablity to the extent that the activity of RNA polymerase is maintained. The mutant RNA polymerase of the present invention prepared by using a thermostable RNA polymerase as the "wild type RNA polymerase" shall be thermostable. As a result, for example, it can be used in PCR to synthesize RNA fragments for sequencing in situ, i.e., during PCR, by using the PCR product as a template.

T7 RNA polymerase has been known to be a promoter specific RNA polymerase with an extremely high specificity. The nucleotide sequence and production method of T7 RNA polymerase are reported in Davanloo et al., Proc. Natl. Acad. Sci. USA., 81:2035-2039 (1984). Its large scale production has been

already described in Zawadzki et al., Nucl. Acids Res., 19:1948 (1991). This phage-derived RNA polymerase can pursue the transcription reaction with a single polypeptide, unlike RNA polymerases of E. coli and higher organisms. (Chamberlin et al., Nature, 228:227-231, 1970). Therefore, it is a particularly excellent material for analyzing the mechanism of transcription, and many mutants have been isolated and reported. Further, the results of its crystallographic analysis are mentioned in Sousa et al., Nature, 364:593-599, 1993.

As other promoter specific RNA polymerases of high specificity, 3 kinds of RNA polymerases derived from T3 phage which infects E. coli, SP6 phage which infects Salmonella, and K11 phage which infects Klebsiella pneumoniae have been well known.

The 4 kinds of RNA polymerases mentioned above are quite resemble to one another in their primary structure of amino acids, sequence of promoter and the like as described hereinafter.

The RNA polymerase of the present invention has an increased ability of incorporating 3'-deoxyribonucleotides and derivatives thereof in comparison with the ability of a corresponding wild type RNA polymerase. As described above, wild type RNA polymerases poorly incorporate 3'-deoxyribonucleotides in comparison with ribonucleotides, which has obstructed their use in nucleotide sequencing. In contrast, the RNA polymerase of the present invention is modified so as to have the ability of incorporating 3'-deoxyribonucleotides and derivatives thereof at least twice higher than that of wild type. The incorporation of 3'-deoxyribonucleotides tends to be decreased especially when 3'-deoxyribonucleotide derivatives are labeled with a fluorescent tag. The RNA polymerase of the present invention can also improve incorporation of such 3'-deoxyribonucleotide derivatives.

The term ribonucleotide herein used means ribonucleoside 5'-triphosphates including ATP, GTP, CTP, UTP

and derivative thereof, and 3'-deoxyribonucleotide means 3'-dATP, 3'-dGTP, 3'-dCTP and 3'-dUTP, and the derivative thereof means, for example, compounds composed of these 3'-deoxyribonucleotides which have a fluorescent label.

The RNA polymerase of the present invention is that at least one of amino acids in a corresponding wild type RNA polymerase is modified. This will be explained in detail hereinafter.

On the basis of the aforementioned various reports about T7 RNA polymerase, the present inventors tried to construct a mutant RNA polymerase which has little or no bias for incorporation efficiency valuable depending on the kind of ribonucleotides observed for T7 RNA polymerase. Various mutants were actually prepared to determine, in particular, which amino acids on wild type RNA polymerases should be mutated, and what kind of amino acids should be used for substitution when substitution is used as mutation. Then, it was found that the ability of incorporating 3'-deoxyribonucleotides and derivatives thereof can be improved by modifying at least one amino acid of wild type RNA polymerases, and completed the mutant RNA polymerase of the present invention.

The present inventors first constructed an expression plasmid pT7R inserted with the T7 RNA polymerase gene, and then mutants of T7 RNA polymerase were constructed based on the expression plasmid pT7R. That is, mutant T7 RNA polymerases, F644Y, F646Y, F667Y, F733Y, F782Y, and F882Y were constructed in which F (phenylalanine) residue of T7 RNA polymerase was replaced with Y (tyrosine) residue, and the ability of incorporation of these mutants was compared. Properties of Y639F mutant of the T7 RNA polymerase, which is a mutant at a location corresponding to Y526 of T7 DNA polymerase, are described in the literature (Sousa., EMBO J., 14:4609-4621 (1995)). Y639F mutant was also constructed, which has a mutation within the residue 631-640, those suggested to change their specificity for dNTP in Japanese Patent Unexamined Publication (KOKAI) No. (Hei) 8-205874/1996.

The amino acid sequence of wild type T7 RNA polymerase mentioned in this specification is based on the sequence encoded by nucleotides 3171-5822 of the T7 phage RNA sequence from the gene sequence database GeneBank, accession No. V01148 J02518 X00411 (39,937 base pairs) (cf. Figures 1 and 2). The upper sequences represented in Figures 1 and 2 are nucleotide sequences, and the lower sequences are amino acid sequences corresponding to the nucleotide sequences. For the nucleotide sequences, the numerals at the right ends are numbers of T7 phage genome registered at GeneBank (Locus T7CG, 39,937 base pairs), and the numerals at the right ends for the amino acids are appended from the first M (methionine) of T7 RNA polymerase starting with 1 and indicate that the full length consists of 883 amino acid residues.

This amino acid sequence is identical to the amino acid sequence reported in Moffatt et al., J. Mol. Biol., 173(2): 265-269, 1984 mentioned above.

Accordingly, the amino acid sequence and the numerals appended to each of the amino acids of wild type T7 RNA polymerase gene in this specification are basically the sequence and numbers represented in Figures 1 and 2. However, as described above, the aforementioned wild type T7 RNA polymerase may contain substitution, insertion and/or deletion which is not the modification intended by the present invention. Therefore, in the case that the wild type RNA polymerase, to which mutation should be introduced for the purpose of the present invention, is a wild type T7 RNA polymerase with other mutation, especially that such mutation is insertion or deletion of amino acids, numbers appended to amino acids are changed due to such insertion and deletion. A wild type T7 RNA polymerase having such insertion and deletion is a member of the wild type T7 RNA polymerase, to which a mutation intended by the present invention should be introduced, so long as it maintains T7 RNA polymerase activity even though its amino acid numbers are different from the numbers represented in Figures 1 and 2.

The amino acid numbers in sequences of RNA polymerases other than T7 RNA polymerase are decided as shown in the sequences listed in Figures 3 and 4. Those may also have substitution, insertion and/or deletion other than the modification intended by the present invention. Accordingly, like the amino acid sequence and the numbers appended to T7 RNA polymerase, when they have such a mutation by insertion or deletion of amino acids, the amino acid numbers are changed due to such insertion and deletion, and a wild type T7 RNA polymerase having such insertion and deletion is a member of the wild type T7 RNA polymerase to which a mutation intended by the present invention should be introduced.

The T7 RNA polymerase gene is prepared as follows: T7 phage DNA is purified. Separately, a primer specific for upstream of N-terminus amino acid region of the T7 RNA polymerase gene (T7Rpol-N: 5'-ATA TTT TAG CCA TGG AGG ATT GAT ATA TGA ACA CGA TTA ACA TCG CTA AG-3') and a primer specific for downstream of C-terminus amino acid region of the same (T7Rpol-C: 5'-ATA TTT TAG CCA TGG TAT AGT GAG TCG TAT TGA TTT GGC G-3') are synthesized. The phage DNA is used as a template for PCR, and thus an expression vector pT7R can be constructed (cf. Example 1). This expression vector can be transformed into E. coli DH5α, and the transformed cells express a large amount of T7 RNA polymerase protein when isopropyl-β-D-thiogalactopyranoside (IPTG) is added.

When the sequence of this T7 RNA polymerase gene prepared as described above was compared with the amino acid sequence shown in Figures 1 and 2, the both sequences completely confirmed to each other. The amino acid sequence shown in Figures 1 and 2 and the amino acid sequence reported in Grachev et al., Bioorg. Kim., 10:824-843, 1984 are different in that the 623rd Y and the 665th L in the amino acid sequence represented in Figures 1 and 2 are replaced with H (623rd) and P (665th) respectively in the amino acid sequence reported by Grachev et al. As described above, wild type RNA polymerases, which are the basis of the mutant RNA polymerase of the present

invention, may contain substitution, insertion, and/or deletion of amino acids with respect to the sequence shown in Figures 1 and 2, which is not the modification intended by the present invention, and the amino acid sequence reported by Grachev et al. where the 623rd and the 665th residues are H and P respectively is included in a member of the wild type RNA polymerases to be a basis of the mutant RNA polymerase of the present invention.

The T7 RNA polymerase purified from E. coli harboring the expression vector pT7R exhibited sufficient RNA synthesis activity in vitro in the presence of DNA containing T7 promoter. Based on this expression plasmid pT7R, the above-mentioned Y639F, F644Y, F646Y, F667Y, F733Y, F782Y, and F882Y were constructed as mutant T7 RNA polymerases, and incorporation ability of these mutants was compared.

For the mutant T7 RNA polymerase having F644Y mutation, another mutation for replacing L665, which is adjacent to F664, with P was introduced in addition to the mutation of F644 according to the report of Grachev et al. mentioned above. That is, mutations of F644Y/L665P were introduced to examine the influence of L665P. Also for the mutant T7 RNA polymerase having F667Y mutation, another mutation for replacing L665, which is adjacent to F667, with P was introduced in addition to the mutation of F667 according to the report of Grachev et al. mentioned above. That is, mutations of F665P/F667Y were introduced.

A mutant T7 RNA polymerase which is introduced with F644Y/L665P/F667Y mutations was also constructed. Comparison of incorporation ability of these mutants was also performed.

The T7 RNA polymerases introduced with mutations were purified, and their abilities of promoter sequence specific RNA synthesis and incorporation of ribonucleoside 5'-triphosphates including ATP, GTP, CTP, UTP and derivatives thereof, as well as 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof were compared with those of wild type T7 RNA polymerase. The results are shown in Table 1 hereinafter.

As a result, as shown in Table 1, F644Y, F644Y/L665P, L665P/F667Y and F644Y/L665P/F667Y maintained sufficient RNA synthesis activity, and showed marked improvement of incorporation of 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof. The incorporation ability of the F644Y/L665P mutant was comparable to that of the F644Y mutant. From these results, it can be seen that the substitution of proline for leucine at 665 do not affect on the incorporation of 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof. While the results are shown only for the L665P/F667Y mutant in Table 1, the F667Y mutant also showed the incorporation ability comparable to that of the L665P/F667Y mutant. The incorporation ability of the F644Y/L665P/F667Y mutant was the highest. While not shown in Table 1, the incorporation ability of the F644Y/F667Y mutant was almost equal to that of the F644Y/L665P/F667Y mutant.

The F782Y mutant maintained RNA synthesis activity, and showed slightly improved ability for incorporating 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof. The F733Y mutant showed slightly decreased RNA synthesis activity, but showed slightly improved ability for incorporating 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof. The F646Y mutant maintained RNA synthesis activity, but showed no improvement of ability for incorporating 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof. The F882Y mutant is not mentioned in Table 1, because it showed markedly decreased RNA synthesis activity.

The Y639F mutant of the T7 RNA polymerase, which has the mutation at a location corresponding to Y526 of T7 DNA polymerase, maintained RNA synthesis activity, but showed no improvement of ability for incorporating 3'-dATP, 3'-dGTP, 3'-dCTP, 3'-dUTP and derivatives thereof.

The results mentioned above suggest that the RNA polymerase of the present invention is particularly an RNA polymerase having modification of at least one of amino acids present in the "nucleotide binding site" of the polymerase and

that such a modification can enhance the ability for incorporating 3'-deoxyribonucleotides and other ribonucleotide analogues in comparison with the ability for corresponding ribonucleotides.

The amino acids present in the above "nucleotide binding site" can be, for example, amino acids in a loop between the helix Y and the helix Z and/or amino acids in a loop between the helix Z and the helix AA of wild type RNA polymerase.

From the steric structure shown in the literature of Sousa et al. (Nature, 364:593-599, 1993), the loop (corresponding to amino acid residues 635 to 647 of T7 RNA polymerase) between the helix Y (corresponding to amino acid residues 625 to 634 of the same) and the helix Z (corresponding to amino acid residues 649 to 658 of the same) and/or the loop (corresponding to amino acid residues 659 to 684 of the same) between the helix Z and the helix AA (corresponding to amino acid residues 685 to 699 of the same), which face the inside of the crafts in the polymerase molecule enclosing template DNA, are considered to constitute a part of the ribonucleotide binding site, which locates quite near the nucleotides. In the present invention, the F residues present at 644, 646 and 667 in a region corresponding to the loops were actually replaced with Y residues (see Figure 5).

The F residues of 733, 782 and 882 are present in a region other than that corresponding to the loop, and considered to face the inside of the crafts in the polymerase molecule. These F residues were also actually replaced with Y residues.

The present invention further relates to an RNA polymerase which has modification at an amino acid selected from those in a region corresponding to the amino acid residues 641-667 of the RNA polymerase derived from T7 phage. The region corresponding to the amino acid residues 641-667 of the RNA polymerase derived from T7 phage correspond to the above-mentioned "nucleotide binding site".

The above-mentioned four RNA polymerases extremely

resemble one another in their primary structures of amino acids, sequence of promoter and the like. In Figures 3 and 4, alignment of amino acid sequences of the aforementioned four RNA polymerases derived from the phages is represented. From this alignment, it can be seen that the RNA polymerases derived from T7, T3, and K11 highly resemble one another. In particular, the amino acid sequences of RNA polymerases derived from T7 and T3 phages show extremely high similarity as shown in Figures 6 and 7. It is conformable to the fact that both of T7 and T3 phages are those infecting E. coli, and they are also resemble each other in their properties. Further, the promoter sequences recognizing these two RNA polymerases also resemble each other, and they have known to have extremely high recognition specificity. Thus, the results obtained in T7 RNA polymerase are relatively readily applied to other RNA polymerases having similar amino acid sequences.

From these high homologies, it can be concluded that a region corresponding to the amino acid residues 644-667 of the RNA polymerase derived from T7 phage in RNA polymerases other than the RNA polymerase derived from T7 phage is the amino acid residues 642-668 for the RNA polymerase derived from T3 phage, the amino acid residues 664-690 for the RNA polymerase derived from K11 phage, and the amino acid residues 633-670 for the RNA polymerase derived from SP6 phage. The RNA polymerases derived from T7, T3, and K11 phages extremely resemble one another as described above, and the results obtained for T7 RNA polymerase can be applied for other RNA polymerases having a similar amino acid sequence (see Figure 8).

As an example of such other RNA polymerases, RNA polymerase derived from K11 phage having tyrosine at the amino acid residue 644 or 667 can be mentioned. RNA polymerase derived from T3 phage having tyrosine at the amino acid residue 645 or 668 can also be exemplified. RNA polymerase derived from K11 phage having tyrosine at one or more of the amino acid residues 664-669 and 690 can further be exemplified. RNA polymerase derived from SP6 phage having tyrosine at one or

more of the amino acid residues 633-638 and 670 can still further be exemplified.

The modification of such an amino acid may be not only substitution of amino acid but also insertion or deletion of amino acid. The mutation of amino acid is, for example, substitution of tyrosine for at least one amino acid residue in a naturally occurring amino acid sequence. The amino acid to be replaced may be, for example, phenylalanine. However, the amino acid to be replaced is not limited to phenylalanine, and any amino acid may be replaced so long as it can enhance the ability for incorporating 3′-deoxyribonucleotides and other ribonucleotide analogues relative to ability for the corresponding ribonucleotides.

Among the mutant RNA polymerases of the present invention, the mutant T7 RNA polymerases F644Y, L665P/F667Y and F644Y/L665P/F667Y maintained sufficient RNA synthesis activity, and showed markedly improved ability for incorporating 3'-dNTPs, and the strong bias observed in the wild type is markedly reduced in these polymerases. Use of T7 RNA polymerase F644Y, L665P/F667Y or F644Y/L665P/F667Y having such characteristics enables a nucleotide sequence determination method utilizing transcription products, which is of more excellent practical applicability in comparison with a nucleotide sequence determination method utilizing a DNA polymerase.

E. coli strains pT7RF644Y (DH5α) and pT7RL665P/F667Y (DH5α), which produce the mutant T7 RNA polymerases F644Y and L665P/F667Y respectively, were already deposited at the National Institute of Bioscience and Human-Technology with international deposition numbers of 5998 (FERM-BP-5998) and 5999 (FERM-BP-5999) respectively on July 2, 1997. E. coli strains pT7RF644Y/L665P/F667Y (DH5α), which produces the mutant T7 RNA polymerase F644Y/L665P/F667Y, was already deposited at the National Institute of Bioscience and Human-Technology with an international deposition number of 6364 (FERM-BP-6364) on May 20, 1998.

The present invention includes a method for producing the aforementioned RNA polymerases of the present invention, which comprises preparing a nucleic acid molecule encoding an RNA polymerase, introducing a mutation into the nucleic acid molecule so that one or more nucleotides in one or more regions should be mutated, and collecting a modified RNA polymerase expressed by the mutated nucleic acid molecule. The preparation of the nucleic acid molecule encoding RNA polymerase, introduction of mutation into the nucleic acid molecule, and collection of the modified RNA polymerase can be performed by using conventional methods.

For example, a mutant T7 RNA polymerase can be constructed by the following method. By using an expression vector inserted with a T7 RNA polymerase gene as template, an expression plasmid comprising a region between the HpaI, and NcoI restriction sites in the C-terminus side of T7 RNA polymerase gene which is introduced with a mutation by PCR is constructed. Subsequently, this expression plasmid can be transformed into E. coli DH5α, which can then produce a large amount of a mutant T7 RNA polymerase protein upon addition of isopropyl-β-D-thiogalactopyranoside (IPTG).

According to the present invention, RNA polymerases can be provided which shows little or no bias of the ability for incorporating ribonucleotides and the like, i.e., solve the problems that incorporation of 3'-deoxyribonucleotide and derivatives thereof are difficult in comparison with corresponding ribonucleotides, and that incorporation of ribonucleotides and 3'-deoxyribonucleotides into a sequence is difference between the nucleotides due to a base group accompanied by the nucleotides.

Further, the use of the RNA polymerase of the present invention enables a method for determining nucleotide sequence more excellent than a method for determining nucleotide sequence utilizing a DNA polymerase without complicated operation. In addition, more quick sequencing of DNA can be realized by using an RNA polymerase of the present invention

having thermostability in PCR, for example, in the method for determining nucleotide sequence of DNA disclosed in WO96/14434.

Examples

The present invention will be explained more in detail with reference to the following examples.

Example 1

Cloning of wild type T7 RNA polymerase gene and construction of expression plasmid

T7 phage harbored in E. coli was prepared as follows. E. coli strain C600 was inoculated in 200 ml of LB culture medium (culture medium prepared by dissolving Bacto tryptone 10g, Bacto yeast extract 5g, and NaCl 5g in 1 liter of water, which was adjusted to pH 7.5, and sterilized in an autoclave). When the cell density reached OD (600 nm) = 1.0, the cells were infected with the phage at a multiplicity of infection of about 2. The OD was determined periodically, and when the OD was sharply decreased, the cell residue was removed by centrifugation. The medium was added with NaCl and polyethylene glycol 6000 to final concentrations of 0.5 M and 10% respectively, stirred sufficiently, and left stand overnight to form precipitates. The precipitates were collected by centrifugation, and suspended in SM buffer (10 mM Tris-HCl, pH 7.5, 10 mM $MgSO_4$, 50 mM NaCl, 0.01% gelatin). This T7 phage concentrate was overlaid on CsCl solution layers carefully overlaid in a centrifugation tube (CsCl solutions having concentrations of 1.267 g/ml, 0.817 g/ml, and 0.705 g/ml from the bottom layer), and centrifuged at 22,000 rpm for 2 hours to form a phage layer. A white band of the phage was carefully separated, and dialyzed against TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA) to remove the CsCl. This phage solution was treated with phenol to denature phage protein to purify genomic DNA of T7 phage.

The T7 RNA polymerase gene corresponds to the

3171st-5822nd base pairs in the 39,937 base pairs of the genome DNA [the total nucleotide sequence of T7 genomic gene had already been reported by Dunn et al. (1983, J. Mol. Biol., 166(4):477-535), but it was slightly corrected (see T7 phage DNA sequence of GeneBank, accession No. V01148 J02518 X00411)]. This genomic DNA was used for PCR as a template, and cloned into an expression vector as follows (see Figure 9). That is, the gene encoding the enzyme was amplified by PCR by using a primer specific for upstream of the N-terminus amino acid region of T7 RNA polymerase gene (T7Rpol-N 5'-ATA TTT TAG CCA TGG AGG ATT GAT ATA TGA ACA CGA TTA ACA TCG CTA AG-3') and a primer specific for downstream of the C-terminus amino acid region of T7 RNA polymerase gene (T7Rpol-C 5'-ATA TTT TAG CCA TGG TAT AGT GAG TCG TAT TGA TTT GCG-3'), each containing NcoI restriction site at the 5'-end. This DNA fragment was digested with NcoI, and separated by electrophoresis on 1% agarose gel, and the band of the objective DNA fragment was cut out from the agarose, and purified by using Gene Pure Kit (Nippon Gene). The DNA fragment was ligated to an expression vector pTrc99a (Pharmacia Biotec) which had been digested with NcoI and dephosphorylated to construct pT7R which expressed T7 RNA polymerase at high level. The plasmid pT7R expressing wild type T7 RNA polymerase was transformed into E. coli DH5α, and the cells resistant to antibiotic ampicillin was cultured. The Trc promoter contained in the expression vector pT7R was driven by adding IPTG to the culture medium. Two hours after the addition of IPTG, the E. coli cells were collected, and the total protein was analyzed by SDS-polyacrylamide gel electrophoresis. As a result, a protein band was detected at a location corresponding to about 99 kDa, which is the molecular weight of T7 RNA polymerase, only when IPTG was added. This protein was further purified by a partially modified version of the previously described method of Zawadzki, V et al. 1991, Nucl. Acids Res., 19:1948 (details may be substantially the same as those of the method for purifying mutant T7 RNA polymerase exemplified in Example 3), and found to have RNA

polymerase activity which was exerted in a T7 promoter specific manner.

Example 2

Construction of expression plasmid for producing mutant T7 RNA polymerases

(1) Construction of expression plasmid for producing mutant T7 RNA polymerase F644Y (see Figure 10)

By using pT7R inserted with the wild type T7 RNA polymerase gene as a template, mutation was introduced by PCR into the region between the HpaI and NcoI restriction sites corresponding to the C-terminus side of the T7 RNA polymerase gene. More precisely, the region was divided into two fragments on the left side and right side of the nucleotide to be mutated, and these DNA fragments were amplified by PCR using primers F646Y(+) (5'-GTT GAC GGA AGC CGT ACT CTT TGG AC-3') introduced with a mutation and F646Y(-) (5'-GTC CAA AGA GTA CGG CTT CCG TCA AC-3'), and primers T7RNAP-HpaI-N (5'-CGC GCG GTT AAC TTG CTT CCT AG-3') and pTrc99a-PstI-C (5'-GCA TGC CTG CAG GTC GAC TCT AG-3'), each containing a restriction cleavage site at the 5'-end. These DNA fragments had complementary regions, and denaturation, annealing and extension reactions of the regions were repeated to prepare a DNA fragment introduced with the desired mutation. This DNA fragment was purified by collecting only a DNA fragment of a desired size through agarose gel electrophoresis, and this was re-amplified by using it as a template together with the primers T7RNAP-HpaI-N and pTrc99a-PstI-C, and cleaved with restriction endonuclease HpaI and PstI. This DNA fragment was separated by 1% agarose gel electrophoresis, and the band of the desired DNA fragment was cut out, and purified. The HpaI-PstI DNA fragment of pT7R was replaced with this DNA fragment to introduce a mutation. The resulting pT7R was transformed into E. coli DH5 α, and cells harboring the plasmid introduced with the mutation were selected. Finally, the nucleotide sequence was determined to confirm whether the

mutation was introduced into the desired site. Thus, the expression plasmid pT7RF644Y for producing mutant T7 RNA polymerase F644Y was obtained. For the production of the mutant T7 RNA polymerase F644Y from this plasmid, expression could be induced by adding IPTG to the cultured E. coli cells harboring the plasmid, like the production of wild type T7 RNA polymerase.

(2) Construction of expression plasmid for producing mutant T7 RNA polymerase L665P/F677Y (see Figures 11 and 12)

The construction of mutant T7 RNA polymerase L665P/F667Y was performed as follows based on PCR technique as in the construction of the F644Y mentioned above.

First, a XhoI restriction site (CTCGAG) was introduced into the T7 RNA polymerase gene region of the expression vector pT7R having the wild type T7 RNA polymerase gene to facilitate the introduction of mutation. More specifically, the expression vector pT7R used as template was amplified by using a primer pair of primer ApaF1 (5'-CAT CTG GTC GCA TTG GGT CAC-3') and primer Xho-R (5'-CCA AGT GTT CTC GAG TGG AGA-3'), and a primer pair of a primer Xho-F (5'-CTA AGT CTC CAC TCG AGA ACA CTT GG-3') and a primer AflII-R (5'-CAG CCA GCA GCT TAG CAG CAG-3'), respectively. The former amplified DNA fragment was digested with restriction endonucleases ApaI and XhoI, and the latter amplified DNA fragment with restriction endonucleases AflII and XhoI, and they were ligated to the expression vector pT7R preliminarily treated with ApaI and AflII by using T4 DNA ligase. This reaction product was transformed into E. coli DH5α, and several colonies grown on an agar plate containing antibiotic ampicillin were obtained. Some of these colonies were selected and cultured, and plasmid DNA was extracted from the cultured cells to obtain plasmid pT7R-Xho in which a XhoI restriction site was introduced in the T7 RNA polymerase gene region (see Figure 10). Presence of this XhoI site can be confirmed by cleavage by a treatment with the restriction endonuclease XhoI, and nucleotide sequencing of the DNA. Using this plasmid pT7R-Xho as a template, PCR was performed

with a primer pair of primer Xho-R and primer 667R (5'-GCT GAG TGT ACA TCG GAC CCT-3'), and a primer pair of a primer 667F (5' of -GCT GAG TGT ACA TCG GAC CCT-3') and a primer AflIIR. The PCR products were directly used as templates for the nucleotide sequencing of the DNA to determine the sequences of the primers 667R and 667F. Then, they were subjected to electrophoresis on 2% agarose gel (Agarose X from Nippon Gene was used as the agarose) respectively, and bands corresponding to DNA fragments of the desired sizes were cut out to purify the DNA fragments by using Gene Pure Kit. The purified two kinds of DNA fragments were mixed, and used as templates for PCR using the primers XhoF and AflIIR. After confirming that the amplified DNA fragment was the desired fragment by restriction mapping and DNA sequencing, the fragment was digested with restriction endonucleases XhoI and AflII, and the resulting fragment was ligated to the plasmid pT7R-Xho preliminarily treated with restriction endonucleases XhoI and AflII by using T4 DNA ligase. This reaction product was transformed into E. coli DH5α, and several colonies of the cells grown on an agar plate containing antibiotic ampicillin were obtained. Some of these colonies were selected and cultured, and plasmid DNA was extracted from the cultured cells. The plasmid DNA was confirmed if it was introduced with the desired mutation by DNA sequencing to finally construct an expression plasmid pT7RL665P/F667Y for producing the mutant T7 RNA polymerase L665P/F667Y (see Figure 12). For the production of the mutant T7 RNA polymerase L665P/F667Y from this plasmid, expression could be induced by adding IPTG to the cultured E. coli cells harboring the plasmid, like the production of wild type T7 RNA polymerase.

Example 3

Purification of mutant T7 RNA polymerases

Mutant T7 RNA polymerase proteins introduced into E. coli were purified.

Wild types of this protein have already been described

in Chamberlin, M et al. Nature, 228:227-231(1970), Davanloo et al., Proc. Natl. Acad. Sci. USA., 81:2035-2039 (1984). Its large scale production has also been reported by Zawadzki, V et al., Nucl. Acids Res., 19:1948 (1991).

All of the mutant T7 RNA polymerases can be purified by principally the same method. The difference of mutation site may cause some difference in the expression level, and behavior in column chromatography. The purification method of mutant T7 RNA polymerase F644Y is exemplified hereinafter. The expression vector pT7RF644Y for F644Y was introduced into E. coli DH5 α, and the cells were cultured in a test tube containing LB culture medium containing antibiotic ampicillin. When the OD (600 nm) of the medium reached 0.4-0.6, isopropyl-β-thiogalactopyranoside (IPTG) was added to the culture to a final concentration of 0.4 mM, and the cultivation was further continued for additional 8 hours. Then, the E. coli cells were collected by centrifugation. Typically, 2 liters of culture medium affords 10 g of E. coli cells in wet weight. If the E. coli cells are not used immediately, they can be stored in a refrigerator at -20℃. Subsequent steps for purification of enzyme should be performed at a temperature lower than room temperature, preferably 0-5℃ unless otherwise indicated. The E. coli cells were washed with 10 times relative to the cell weight of a washing buffer (20 mM Tris-HCl, pH 8.1, 130 mM NaCl, 2 mM $EDTANa_2$ at 25℃), centrifuged again (5,000 x g, 4℃, 10 minutes), suspended in 10 times in volume of a sonication buffer [50 mM Tris-HCl, pH 8.1, 100 mM NaCl, 0.1 mM $EDTANa_2$, 5 mM dithiothreitol (DTT), 0.1 mM benzamidine, 30 μg/ml phenylmethylsulfonyl fluoride (PMSF), 10 μg/ml bacitracin], and sonicated by using Sonifier 450 (Branson) at 80W for more than 15 minutes to destroy the cells and reduce the viscosity of the cells. Then, the cell suspension is centrifuged at 12,000 x g and 4℃ for ten minutes to remove the cell debris. 10% streptomycin sulfate was slowly added dropwise to the resulting supernatant to a final concentration of 2.0% with stirring, and stirring was further continued for 30 minutes.

The supernatant was centrifuged at 12,000 x g and 4℃ for ten minutes to remove precipitates, and slowly added with ammonium sulfate powder with stirring to form precipitates. In this case, precipitates were first collected by 30% saturation of ammonium sulfate (30% ammonium sulfate precipitation), and the resulting supernatant was further added with ammonium sulfate to 60% saturation with stirring to form precipitates again (30-60% ammonium sulfate precipitation). The supernatant was added again with ammonium sulfate powder to 90% ammonium sulfate saturation, and stirred at 4℃ for 1 hour, and the precipitates were collected by centrifugation. Aliquots of these three ammonium sulfate fractions were analyzed for proteins by SDS-acrylamide gel electrophoresis, and it was found that most of the objective mutant T7 RNA polymerase was present in the 30-60% ammonium sulfate fraction. Therefore, purification was performed hereafter by using this fraction. The.30-60% ammonium sulfate fraction was suspended in a small amount of column buffer (20 mM KPO_4, pH 7.7, 100 mM NaCl, 1mM DTT, 30 μg/ml PMSF), and desalted by dialysis against 500 ml of the same buffer for 16 hours. The dialysate was applied on a heparin-Sepharose column of 5 ml volume (Pharmacia Biotec). Subsequently, the column was washed with the same buffer until any material absorbing ultraviolet ray at 280 nm disappeared, and eluted with a linear gradient of 0.1 M to 0.64 M NaCl in the same buffer of about 40 times volume of the column volume. The eluent was collected in test tubes as fractions of a suitable volume, and immediately subjected to SDS-acrylamide gel electrophoresis for protein analysis to identify fractions containing proteins around a molecular weight considered to be of the objective T7 RNA polymerase. In typical examples, it should be found around 0.4 M NaCl. The fractions containing the protein were collected, and desalted by dialysis against about 1 liter of the column buffer (20 mM KPO_4, pH 7.7, 100 mM NaCl, 1 mM DTT, 30 μg/ml PMSF) for 16 hours. The fractions desalted by dialysis were applied to a Q-Sepharose column (Pharmacia Biotec) of 5 ml volume that preliminarily

equilibrated with the same buffer, and the column was washed with the same buffer until any material absorbing ultraviolet ray at 280 nm disappeared, and eluted with a linear gradient of 0.1 M to 0.64 M NaCl in the same buffer of about 40 times volume of the column volume. The eluent was collected in test tubes as fractions of a suitable volume, and immediately subjected to SDS-acrylamide gel electrophoresis for protein analysis to identify fractions containing proteins around a molecular weight considered to be of the objective T7 RNA polymerase. In typical examples, it should be found around 0.24 M NaCl. The fractions containing the protein were collected, dialyzed against 500 ml of storage buffer (50% glycerol, 20 mM KPO_4, pH 7.7, 100 mM NaCl, 1 mM DTT, 30 μg/ml PMSF) for 16 hours, and stored at -20℃ until use. In vitro RNA synthesis activity and activity of the contaminated ribonuclease of this sample were examined. The in vitro RNA synthesis activity was examined by, for example, performing RNA synthesis reaction according to the enzyme dilution method by using the plasmid containing T7 promoter as a template and a commercially available wild type T7 RNA polymerase (BRL, Gibco) as a standard, and subjecting the synthesized RNA to agarose gel electrophoresis to estimate approximate titer. In this case, because degree of decomposition of RNA is also determined, simple assay for contaminated ribonuclease can simultaneously be performed. As a typical example, 2,500,000 units of the mutant T7 RNA polymerase F644Y protein was purified from 1 liter of culture medium using the above-described steps, and this preparation was substantially free from RNase contamination.

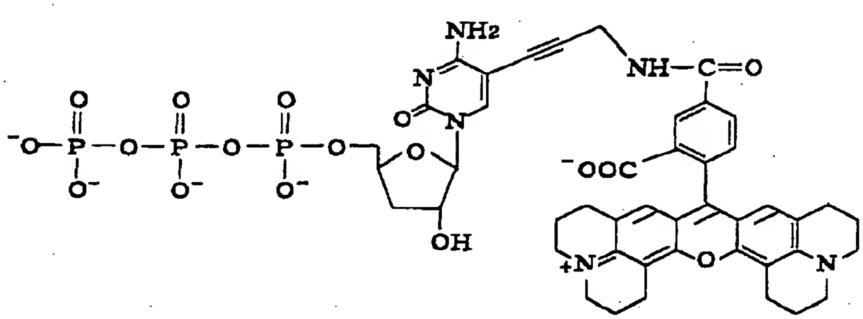
Example 4

Improvement of incorporation ratio of 3'-dNTP derivatives

3'-dNTP incorporation efficiency of the purified mutant T7 RNA polymerases F644Y and L665P/F667Y was compared with that of wild type T7 RNA polymerase as follows. In vitro transcription reaction was performed by, for example, a

partially modified version of the method of Melton, D.A, [Nucleic Acids Res., 12: 7035-7056 (1984)]. More specifically, the reaction was performed in a total volume of 10 μl containing a plasmid vector pBluescriptKS(+) having T7 promoter (Stratagene) linealized by the reaction with a restriction endonuclease PvuII or ScaI as a template, 150 μM of 5-carboxy-X-rhodamine-labeled 3'-deoxycytidine-5'-triphosphate which was a dye terminator prepared according to the method described in WO96/14434 as a derivative of 3'-dNTP, 500 μM of GTP and UTP, 250 μM of ATP and CTP, 8 mM of $MgCl_2$, 2 mM of spermidine-$(HCl)_3$, 5 mM of DTT, 40 mM of Tris/HCl pH 8.0 (BRL, Gibco) and 25 units of wild type T7 RNA polymerase (BRL, Gibco or Nippon Gene) or the mutant T7 RNA polymerase F644Y or L665P/F667Y at 37℃ for 1hour. Then, to remove the unreacted dye terminator remained in the reaction product, the transcription product was purified by gel filtration using Sephadex G-50 column (Pharmacia Biotec), and the purification product was evaporated to dryness using a centrifugal evaporator.

The above 5-carboxy-X-rhodamine-labeled 3'-deoxycytidine-5'-triphosphate is a compound represented by the following chemical formula:



The dried reaction product was dissolved in 6 μl of formamide/EDTA/Blue dextran loading buffer according to the instruction manual Ver.1.0 of ABI PRISM 377 DNA Sequencing

System available from Perkin-Elmer Japan, and 2 μl of the solution was analyzed by ABI 377 DNA Sequencer and an analysis program using denatured gel for sequencing analysis which contained 6M urea/4% Long Ranger™ acrylamide solution (FMC). The results are shown in Figure 13 as a gel image. It was found that the mutant T7 RNA polymerase F644Y could afford a sequence ladder 3 times longer than that afforded by the wild type T7 RNA polymerase, and a transcription product of about 700 bases was also confirmed.

The peak intensities of the sequence ladders obtained by using F644Y and L665P/F667Y are shown in Figure 14 and Figure 15 with the peak intensity obtained by using wild type T7 RNA polymerase. From this comparison, it was confirmed that altitude of the peaks for the mutant enzymes showed less fluctuation in comparison with the wild type, and the peak showed stronger signals. This indicates that the mutation of F644Y or L665P/F667Y improved the incorporation efficiency for 3'-dCTP derivatives for this case, and that transcription reaction by these mutant T7 RNA polymerases exhibits ladder extension characteristic comparable to the data productivity of the conventional methods for determining nucleotide sequence using a DNA polymerase.

Example 5

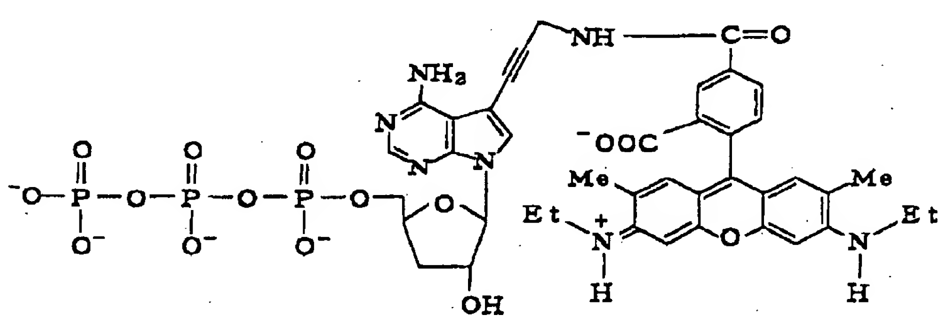
Example of sequencing reaction by the dye terminator method utilizing mutant T7 RNA polymerase

Sequencing reaction by the dye terminator method was performed utilizing the purified mutant T7 RNA polymerases F644Y and L665P/F667Y, and the wild type T7 RNA polymerase as follows for comparison.

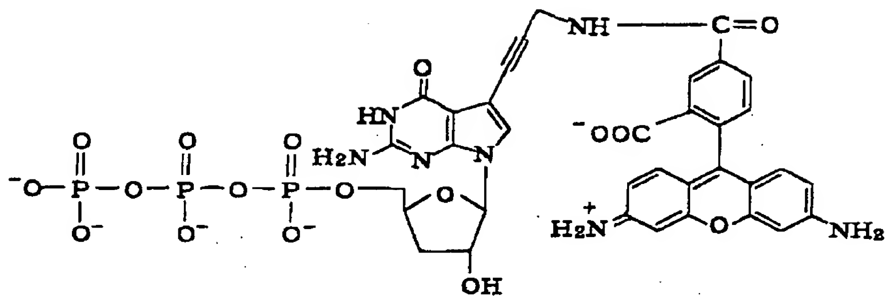
For the in vitro transcription reaction, the method of Melton, D.A. (1984, Nucleic Acids Res., 12:7035-7056) exemplified in Example 4 was used. More specifically, the reaction was performed in a total reaction volume of 10 μl containing a plasmid vector pBluescriptKS(+) having T7 promoter linealized by the reaction with a restriction

endonuclease PvuII or ScaI as a template, 5-carboxyrhodamine 6G-labeled 3'-deoxyadenosine-5'-triphosphate, 5-carboxyrhodamine 110-labeled 3'-deoxyguanosine-5'-triphosphate, 5-carboxy-X-rhodamine-labeled 3'-deoxycytidine-5'-triphosphate, and 5-carboxytetramethylrhodamine-labeled 3'-deoxyuridine-5'-triphosphate, which were dye terminators prepared according to the method described in WO96/14434 as derivatives of 3'-dNTP, 500 μM of GTP and UTP, 250 μM of ATP and CTP, 8 mM of $MgCl_2$, 2 mM of spermidine-$(HCl)_3$, 5 mM of DTT, 40 mM of Tris/HCl pH 8.0 (BRL, Gibco) and 25 units of wild type T7 RNA polymerase (BRL, Gibco or Nippon Gene) or the mutant T7 RNA polymerase F644Y at 37℃ for 1hour. Then, to remove the unreacted dye terminators remained in the reaction product, the transcription product was purified by gel filtration using Sephadex G-50 column (Pharmacia Biotec), and the purification product was evaporated to dryness using a centrifugal evaporator.

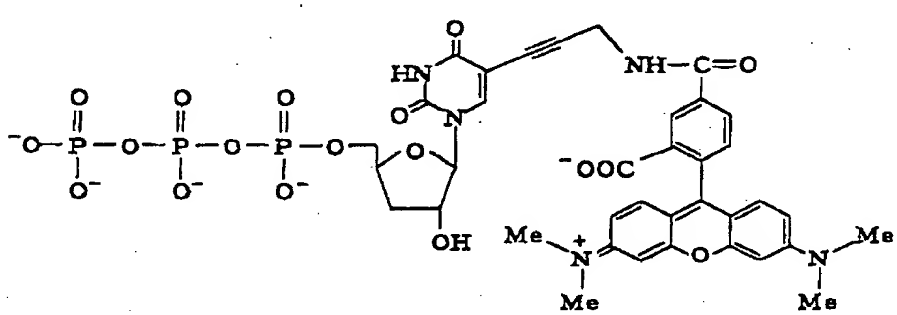
The above 5-carboxy-X-rhodamine-labeled 3'-deoxycytidine-5'-triphosphate is the same compound as used in Example 4. 5-Carboxyrhodamine 6G-labeled 3'-deoxyadenone-5'-triphosphate, 5-carboxyrhodamine 110-labeled 3'-deoxyguanosine-5'-triphosphate, and 5-carboxytetramethylrhodamine-labeled 3'-deoxyuridine-5'-triphosphate are the compounds represented by the following chemical formulae:



5-Carboxyrhodamine 6G-labeled 3'-deoxyadenone-5'-triphosphate



5-Carboxyrhodamine 110-labeled 3'-deoxyguanosine-5'-triphosphate



5-Carboxytetramethylrhodamine-labeled 3'-deoxyuridine-5'-triphosphate

The dried reaction product was dissolved in 6 μl of formamide/EDTA/Blue dextran loading buffer according to the instruction manual Ver.1.0 of ABI PRISM 377 DNA Sequencing System available from Perkin-Elmer Japan, and 2 μl of the solution was analyzed by ABI 377 DNA Sequencer and an analysis

program using denatured gel for sequencing analysis which contained 6M urea/4% Long Ranger™ acrylamide solution (FMC). As a result, it was found that the mutant T7 RNA polymerases F644Y and L665P/F667Y could afford higher peak intensity with less fluctuation in comparison with the wild type T7 RNA polymerase, and their sequence reading was possible. When the wild type T7 RNA polymerase was used, its sequence reading was almost impossible.

Example 6

Construction of expression plasmid for producing mutant T7 RNA polymerase F644Y/L665P/F667Y (see Figure 17)

Construction of the mutant T7 RNA polymerase F644Y/L665P/F667Y was performed based on PCR, as in the construction method of the expression plasmid for producing the mutant T7 RNA polymerase L665P/F667Y previously constructed (see Example 2), as follows.

PCR was performed by using the expression plasmid producing the mutant T7 RNA polymerase L665P/F667Y as template together with a primer pair of the primer Xho-F and the primer T7-DOUBLE-R (21-mer: 5'-CTCTTTGGACCCGTAAGCCAG-3') or a primer pair of the primer T7-DOUBLE-F (29-mer: 5'-TTACGGGTCCAAAGAGTACGGCTTCCGTC-3') and the primer AflII-R. The PCR products were directly used as templates and determined for DNA sequences to confirm the sequences of the primers T7-DOUBLE-R and T7-DOUBLE-F. Each of the products was subjected to electrophoresis on 2% agarose gel to purify DNA fragment of the intended size. The purified two kinds of DNA fragments were mixed, and used as template for PCR using the primers XhoF and AflIIR. After confirming that the amplified DNA fragment was the desired fragments by restriction mapping and DNA sequencing, the fragment was digested with restriction endonucleases XhoI and AflII, and the resulting fragment was ligated to the plasmid pT7RL665P/F667Y preliminarily treated with restriction endonucleases XhoI and AflII by using T4 DNA ligase. This reaction product was transformed into E. coli

DH5α, and several colonies of the cells grown on an agar plate containing antibiotic ampicillin were obtained. Some of these colonies were selected and cultured, and plasmid DNA was extracted from the cultured cells. The nucleotide sequence of the plasmid DNA was sequenced to confirm that the desired mutation should be introduced, and thus an expression plasmid pT7RF644Y/L665P/F667Y for producing the mutant T7 RNA polymerase F644Y/L665P/F667Y was finally constructed (see Figure 17). For the production of the mutant T7 RNA polymerase F644Y/L665P/F667Y from this plasmid, expression could be induced by adding IPTG to cultured E. coli cells harboring the plasmid, like the production of the wild type T7 RNA polymerase.

Example 7

Purification of mutant T7 RNA polymerase F644Y/L665P/F667Y

The mutant T7 RNA polymerase F644Y/L665P/F667Y could be purified by the same method as in Example 3. In a typical example, 1,000,000 units of the mutant T7 RNA polymerase F644Y/L665P/F667Y protein was purified from 1 liter of culture medium. The obtained RNA polymerase was detected substantially as a single band, and RNase was not detected in this specimen by SDS-polyacrylamide gel electrophoresis.

Example 8

Improvement of incorporation rate of 3'-dNTP derivatives

Ribonucleotide (NTP) and 3'-deoxynucleotide (3'-dNTP) incorporation rates of the mutant T7 RNA polymerase purified in Example 7 were measured as follows.

pBluescript(KS+) plasmid (Stratagene) linearized by reaction with a restriction endonuclease, PvuII, was used as a template for the transcription reaction, and 250 μM each of ATP, CTP, GTP, and UTP, 2 mM of spermidine-$(HCl)_3$, 5 mM of DTT, 40 mM Tris/HCl pH 8.0, 0.1 μl of [α-^{32}P]UTP (3000 Ci/mmole), and 25 units of the mutant T7 RNA polymerase F644Y/L665P/F667Y were also used for the reaction. For two kinds of reaction mixture (with or without 3'-dATP, final concentration was 100

μM), the reaction was performed at 37℃ for 60 minutes. The whole reaction mixture was spotted on DE81 paper (Whatman), washed three times with phosphate buffer, and dried. The DE81 paper was placed into a scintillation vial, and radioactivity was measured using a scintillation counter (Beckman) for each reaction. Degree of inhibition of the [α-^{32}P]UTP incorporation was calculated by comparing the values obtained with and without 3'-dATP based on the measured radioactivity. The relative activity obtained from calculated inhibition degree and defined as a relative value to the inhibition degree of the wild type T7 RNA polymerase normalized to 1.000 was shown in Table 1.

The inhibition degree was calculated by using the wild type T7 RNA polymerase, T7 RNA polymerase F644Y, L665P/F667Y obtained in Example 3, mutant T7 RNA polymerase F644Y/L665P, F782Y, F733Y, F646Y or Y639F constructed and purified in the same manner as in Examples 2 and 3 for the reaction instead of the above F644Y/L665P/F667Y mutant, and relative activities are shown in Table 1.

In the results of Table 1, a larger value indicates that the corresponding mutant enzyme has a mutation making 3'-dATP incorporation easier in a higher degree. For example, it is meant that the mutant T7 RNA polymerase F644Y/L665P/F667Y is 5.58 times more likely to incorporate 3'-dATP in comparison with the wild type enzyme. It is demonstrate that the F644Y/L665P/F667Y mutant was the mutant enzyme exhibiting the least bias for the 3'-dATP incorporation among the mutant enzymes prepared.

Table 1

Mutation site	Relative activity of RNA polymerase for 3'-dATP
F644Y	5.130
F644Y/L665P	5.130
L665P/F667Y	4.711
F644Y/L665P/F667Y	5.580
F782Y	1.173
F733Y	1.075
F646Y	0.459
Y639F	0.930
Wild type	1.000

Example 9

Example of sequencing reaction utilizing mutant T7 RNA polymerase F644Y/L665P/F667Y

A template used as a template for sequencing reaction was prepared by PCR as follows.

As the template for PCR, human thyroid-stimulating hormone (hTSH-β) cDNA subcloned into a plasmid derived from BS750 having T7 promoter was used. By using this plasmid 100fg having hTSH-β with L220 primer (5'-TAA CAA TTT CAC ACA GGA AAC A-3') and 1211 primer (5'-ACG TTG TAA AAC GAC GGC CAG T-3') existing at the both sides of the cloning site, PCR reaction was performed in a reaction volume of 20 μl (1 cycle of 94℃ for 2 minutes, 30 cycles of 94℃ for 1 minute, 55℃ for 1 minute, and 72℃ for 1.5 minutes, followed by 72℃ for 5 minutes). The T7 promoter existed in the downstream of 1211 primer of the PCR product obtained from the above PCR reaction.

The transcriptional sequencing reaction was performed by the method of Melton, D.A, [Nucleic Acids Res., 12: 7035-7056 (1984)].

1 μl (about 10 ng) of the above PCR product was used for the sequencing reaction. The reaction was performed in a total reaction volume of 10 μl containing the same dye terminators as used in Example 5, 4 μM R6G-3'-dATP [5-carboxyrhodamine 6G-labeled 3'-deoxyadenosine-5-triphosphate (n=4)], 4 μM R110-3'-dGTP [5-carboxyrhodamine

110-labeled 3'-deoxyguanosine-5-triphosphate (n=4)], 80 μM XR-3'-dCTP [5-carboxy-X-rhodamine-labeled 3'-deoxycytidine-5-triphosphate (n=4)], 20 μM TMR-3'-dUTP [5-carboxytetramethylrhodamine-labeled 3'-deoxyuridine-5-triphosphate (n=4)], 500 μM UTP, 250 μM ATP, 200 μM CTP, 500 μM GTP, 2 mM spermidine-$(HCl)_3$, 5 mM DTT, 40 mM Tris/HCl pH 8.0 (BRL, Gibco) and 25 units of the mutant T7 RNA polymerase F644Y/L665P/F667Y at 37℃ for 1hour.

Then, to remove the unreacted dye terminator remained in the reaction product, the transcription product was purified by gel filtration using Sephadex G-50 column (Pharmacia Biotec), and the purification product was evaporated to dryness using a centrifugal evaporator.

The dried reaction product was dissolved in 6 μl of formamide/EDTA/Blue dextran loading buffer according to the instruction manual Ver.1.0 of ABI PRISM 377 DNA Sequencing System available from Perkin-Elmer Japan, and 2 μl of the solution was analyzed by ABI 377 DNA Sequencer and an analysis program (Sequencing Analysis Ver. 3.0) using denatured gel for sequencing analysis which contained 6M urea/4% Long Ranger™ acrylamide solution (FMC) to afford an electropherogram. The results are shown in Figure 18. Excellent sequencing analysis was possible as demonstrated.